ANL/EES-TM-315

AN EVALUATION OF MARKET PENETRATION FORECASTING METHODOLOGIES FOR NEW RESIDENTIAL AND COMMERCIAL ENERGY TECHNOLOGIES
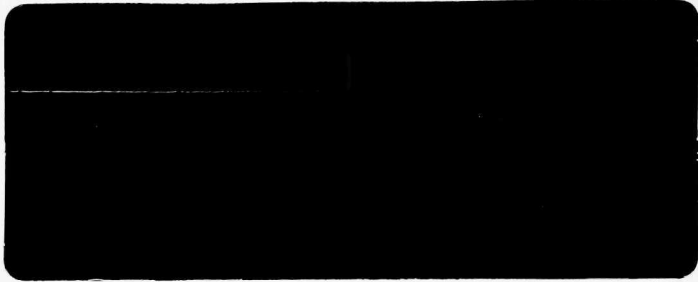
ARGONNE NATIONAL LABORATORY

Energy and Environmental Systems Division

Operated by

THE UNIVERSITY OF CHICAGO for U. S. DEPARTMENT OF ENERGY

under Contract W-31-109-Eng-38

Argonne National Laboratory, with facilities in the states of Illinois and Idaho, is owned by the United States government, and operated by The University of Chicago under the provisions of a contract with the Department of Energy.

ARGONNE NATIONAL LABORATORY
9700 South Cass Avenue, Argonne, Illinois 60439

ANL/EES-TM-315

AN EVALUATION OF MARKET PENETRATION FORECASTING METHODOLOGIES FOR NEW RESIDENTIAL AND COMMERCIAL ENERGY TECHNOLOGIES

by

P.S. Raju* and A.P.S. Teotia

Energy and Environmental Systems Division

May 1985

*University of Illinois at Chicago

CONTENTS

FIGURES

FOREWORD

In 1979, the U.S. Congress created the Office of Minority Economic Impact (OMEI) within the U.S. Department of Energy (DOE) out of concern for the effects of energy shortages and rising prices on minority citizens, particularly those with low incomes. The legislation (42 U.S.C., Sec. 7141[c]) defines a minority group as one consisting of black, Oriental, American Indian, Eskimo, or Aleut citizens, or Puerto Rican or other Spanish-speaking citizens of Spanish descent. This law requires OMEI, among other things, to conduct research to (1) determine the average energy consumption and use patterns of minority groups relative to other population groups and (2) evaluate the percentage of disposable income spent on energy by minority groups relative to other population groups.

As part of its compliance with this mandate, OMEI commissioned Argonne National Laboratory (ANL) to conduct a multiyear research program to determine energy consumption and expenditures by minority groups. The ANL program consists of three tasks:

- Assembling a data base and developing the tools to assess the effects of government energy policies and programs on minority groups.

- Assessing the effects of government programs on minorities and identify options for modifying those programs (e.g., through policy, regulatory, or legislative changes) to alleviate possible hardships for minority groups.

- Providing market research assistance to energy-related businesses owned by members of minority groups.

Further information on the overall OMEI research program can be obtained from Georgia Johnson, the research project officer for the DOE Office of Minority Economic Impact, or from David Poyer, the principal project investigator at ANL. Information related to this report may be obtained directly from the authors.

ACKNOWLEDGMENTS

AN EVALUATION OF MARKET PENETRATION FORECASTING METHODOLOGIES FOR NEW RESIDENTIAL AND COMMERCIAL ENERGY TECHNOLOGIES

by

P.S. Raju and A.P.S. Teotia

SUMMARY

Forecasting market penetration is an essential step in the development and assessment of new technologies. This report reviews several methodologies that are available for market penetration forecasting. The primary objective of this report is to help entrepreneurs understand these methodologies and aid in the selection of one or more of them for application to a particular new technology. This report also illustrates the application of these methodologies, using examples of new technologies, such as the heat pump, drawn from the residential and commercial sector. The report concludes with a brief discussion of some considerations in selecting a forecasting methodology for a particular situation. It must be emphasized that the objective of this report is not to construct a specific market penetration model for new technologies but only to provide a comparative evaluation of methodologies that would be useful to an entrepreneur who is unfamiliar with the range of techniques available.

The specific methodologies considered in this report are as follows: subjective estimation methods, market surveys, historical analogy models, time series models, econometric models, diffusion models, economic cost models, and discrete choice models.

In addition to these individual methodologies, which range from the very simple to the very complex, two combination approaches are also briefly discussed: (1) the economic cost model combined with the diffusion model and (2) the discrete choice model combined with the diffusion model. This discussion of combination methodologies is not meant to be exhaustive. Rather, it is intended merely to show that many methodologies often can complement each other. A combination of two or more different approaches may be better than a single methodology alone.

1 INTRODUCTION

This report surveys the methods available to forecast the market penetration of new technologies. Because the term *new technologies* is rather broad, it is defined in this report to refer to major technological innovations and not marginal improvements over existing technologies or products, although many of the forecasting methods surveyed may apply to the latter category as well. Examples of major technological innovations are new space heating and cooling technologies such as advanced heat pumps, packaged cogeneration systems, and photovoltaics.

Forecasting the market penetration of new technologies is not without difficulty. Many of the sophisticated computer-based models that have been developed for assessing the market for nondurable consumer items (see, for example, Refs. 1 and 2) are not applicable to new technologies for various reasons. First, new technologies are often difficult for the average consumer to comprehend, and they may also entail a considerable change in consumer behavior or lifestyle. The collection of primary data from consumers, therefore, becomes difficult. Second, techniques such as concept testing and market testing are either not applicable or of limited use for new technologies. The lack of primary data, therefore, makes it necessary to rely on secondary data, especially historical data if available.

Good forecasts are important in business planning, and their value to entrepreneurs cannot be overemphasized. Given the importance of forecasting, the specific objectives of this report are to:

- Review the major methodologies available to forecast the market penetration of technological innovations,

- Illustrate the application of these methodologies, and

- Outline some considerations in selecting a forecasting methodology for a particular situation.

To accomplish these objectives, the report describes the major characteristics, advantages, and limitations of each methodology. Because repeat purchase is irrelevant to the market penetration of technologies, the methodologies are primarily concerned with first purchase. Also, certain combinations of methodologies are described to illustrate that individual methods can often be used in a complementary manner for a particular forecasting task. Examples of applying the methodologies are drawn predominantly from the residential and commercial sector and concern such technologies as the advanced heat pump. However, this report does not construct a specific penetration model for new technologies. For heat pumps, selection of a methodology and development of a market penetration model will be addressed in a future report.

The task described in this report is the first step toward preparing a technical information document to help minority entrepreneurs assess the market potential of new energy technologies. Market penetration models could be used to estimate the demand

for new technologies for several years into the future. This, in turn, would assist entrepreneurs in drawing up detailed marketing plans for new technologies. In addition, market projections could be developed for key segments, such as regions, commercial and residential sectors, single-family and multifamily homes, and specific population types (e.g., black, white, poor, elderly, or Hispanic). Such breakdowns would help entrepreneurs to better target their marketing efforts in their particular regions or markets.

2 MARKET PENETRATION FORECASTING METHODOLOGIES

The following methods for forecasting market penetration are considered in this report:

- Subjective estimation methods,

- Market surveys,

- Historical analogy models,

- Time series models,

- Econometric models,

- Diffusion models and technology substitution models,

- Economic cost models, and

- Discrete choice models.

Each of these is discussed below. Section 3 discusses some combinations of these methodologies that are applicable to new technologies.

2.1 SUBJECTIVE ESTIMATION METHODS

Subjective estimates of market penetration over time can be made by top management personnel, experts, or others who are knowledgeable about a new technology. Techniques range from very informal ones such as brainstorming to more formal approaches such as the morphological and Delphi methods. A discussion of these latter techniques is beyond the scope of this report but can be found in other sources.[3,4]

Subjective estimation is particularly useful when a technology is extremely new and no historical data are available. It is also particularly suited to long-range forecasting when experts may only have a "feel" for the market and for changes or discoveries that are likely to occur in the future. (In many markets, sellers know the needs of their buyers so well that they have an intuitive but highly accurate feeling for the market future trends.) The disadvantage of subjective estimation is that it may be expensive relative to the reliability it provides. Forecasts can also be biased by the intent or viewpoint of the forecaster. For example, a utility's forecast for residential photovoltaics may differ dramatically from that made by solar enthusiasts. Even so, subjective forecasting methods can provide a range of estimates about the commercial success of a new technology. Also, no matter how sophisticated the forecasting method used, there always remains an element of residual uncertainty, which makes some subjective judgment almost always necessary. No method can be perfectly accurate, and in that sense, subjective methods complement other forecasting methods.

2.2 MARKET SURVEYS

The survey is a common approach to forecasting market demand, and intention surveys are quite popular for durable consumer items such as automobiles and major appliances. This method could theoretically be used for new technologies as well. For example, for technologies such as the heat pump, either consumers could be surveyed about their attitudes and intentions regarding heat pumps, or entrepreneurs could be surveyed as to the likelihood of consumers adopting heat pumps. Such attitudes and intentions could also be elicited through techniques such as focus groups.

However, several problems exist with the survey method. First, the resulting forecast is based on expected future actions that are subject to change. Intentions are sometimes very poor predictors of actual behavior. Second, the respondents may not be knowledgeable enough to comprehend a new technology or assess its impact even on themselves. Third, even if attitudes and intentions have good predictive power, the forecast would be valid only in the short term. Finally, surveys involve collection of primary data. In cases where a considerable amount of historical data is available, the collection of primary data may not add significantly to the accuracy of the forecast.

2.3 HISTORICAL ANALOGY MODELS

Historical analogy is a forecasting method based on comparison of a new product or technology with an existing one that is analogous. The market penetration share over time is assumed to be essentially the same for both technologies. The method thus does not explicitly consider variables other than time and assumes that the new and the analogous products are similar and exist in identical competitive markets.

Historical analogy is useful especially when historical data are not available, e.g., because the product or technology has been introduced very recently. It can also provide a useful first approximation before more-sophisticated forecasting methods are used. Its success, however, depends on identification of an analogous product or technology and is therefore less likely to be useful for pioneering technologies. Because the method does not consider exogenous variables explicitly, its explanatory value is limited. It could, however, have good predictive value and also has a major advantage in that the data requirements are minimal.

As an example, if this method were used to estimate the market penetration of advanced heat pumps, an appropriate analogous technology must first be identified. Ideally, the analogous technology should have faced competitive and macroeconomic conditions similar to those faced by the heat pump. In one study, Gordian Associates used central air conditioning and electric heat pumps as analogous technologies to forecast penetration of advanced electric and gas heat pumps.[5] However, historical analogy appears to be a weak modeling approach in situations where a considerable amount of historical data, including data on cost relative to competing products or technologies is available. In such cases, it should be possible to use more-sophisticated forecasting methodologies such as those discussed below.

2.4 TIME SERIES MODELS

Time series models constitute one of the oldest methods of forecasting, and the models range from relatively simple to fairly complex and sophisticated. Basically, the models forecast product sales as a function of time and past sales alone. No other explanatory variables are included. There are three classes of time series models:

- Adaptive smoothing,

- Trend extrapolation, and

- Stochastic time series.

2.4.1 Adaptive Smoothing Models

Adaptive smoothing models involve the application of a set of differential weightings to the time series of past observations. The general specification of such models can be written as:

$$y_{t+1} = w_t \, y_t + w_{t-1} \, y_{t-1} + w_{t-2} \, y_{t-2} + \cdots$$

where w is the weight applied to past observations.

Different techniques are available for generating the weighting coefficients.[6] One is the moving-average model, in which the simple average of a certain number of past observations is used to develop the forecast for the next period. It may be unrealistic, however, to assume that a good forecast for the next period would be given by a simple average of all of the past values. It may be more reasonable to expect that values from the recent past will have more effect than earlier values. In such a case, an exponentially weighted moving average (EWMA) model can be used, which is formulated as follows:

$$y_{t+1} = \ell \, y_t + \ell \, (1-\ell) \, y_{t-1} + \ell \, (1-\ell)^2 \, y_{t-2} + \cdots$$

where ℓ is a number between 0 and 1 that indicates how heavily the recent values are weighted compared to older values.

The basic exponential smoothing model is applicable when the demand process is stationary, i.e., while demand may fluctuate from one time period to the next, the average value is reasonably stable. If the average demand series has an upward or downward trend, the EWMA model will underpredict or overpredict future values, respectively. Thus, any trend in the data should be removed before the technique is used. Once an untrended forecast is made, the trend can be reintroduced to obtain a final forecast.

2.4.2 Trend Extrapolation Models

In trend extrapolation models, it is assumed that sales growth or decline in the future will be based on past patterns. Historical data are analyzed to identify the pattern. Trend extrapolation can be quite accurate in the short-to-medium time range if a large historical data base is available. Accuracy declines as forecasts extend farther into the future. A number of trend curves are available to forecast future demand. The most commonly used are linear, exponential, modified exponential, Gompertz, logistic, and generalized logistic functions.[7] The parameters of these functions are estimated from past data using the least-squares method. The selection of the correct trend extrapolation function is based on common sense and on trial and error. The choice can often be made by observing other similar products or technologies or by experimenting with different functions to see which one provides the best fit. Several "rules of thumb" are also available to select the correct function.[8] Once the function is identified, sensitivity analysis might be useful to determine how sensitive the dependent variable is to the parameters of the function.

2.4.3 Stochastic Time Series Models

The principle of stochastic models can be described as follows. Assume that each value in the series y_1, y_2, y_3, $\cdots$ y_t is drawn randomly from a probability distribution. If the characteristics of this randomness are described, it is possible to infer the probabilities associated with alternative future values of the series. In moving average models, each observation y_t is generated by a weighted average of past random disturbances. In autoregressive models, the current observation y_t is generated by past observations, together with a random disturbance for the current period. The autoregressive moving average (ARMA) model combines these two approaches. All these models assume a "stationary" time series, i.e., stochastic processes are assumed to be in equilibrium over time about a constant mean level. The integrated autoregressive moving average (ARIMA) model can be used for homogeneous nonstationary processes, i.e., nonstationary series that can be transformed into stationary series by differencing one or more times. The Box-Jenkins approach is the most commonly used method for implementing a stochastic time series model.[9]

2.4.4 Evaluation of Time Series Models

The main advantage of time series models is that the methodology is well established and therefore easy to apply. These models are only applicable, however, when historical data are available for several years. Hence, the methodology would be unsuitable for technologies in existence for a relatively short time. Also, the validity of the methodology is higher when past trends are likely to continue unchanged. Among the three models discussed above, the adaptive smoothing and trend extrapolation methods are deterministic models that do not model the stochastic component. Since the stochastic component is the main source of error in forecasts, stochastic time series models are superior in forecast accuracy. Because explanatory variables are not included in time series models, their explanatory power is limited although their predictive power may be high.

For the heat pump, considerable historical data are available and hence time series models are a definite option. The Air Conditioning and Refrigeration Institute (ARI), for example, has annual shipment data on heat pumps from 1957 onward that are suitable for time series modeling.[10]

2.5 ECONOMETRIC MODELS

Econometric models use historical data to estimate the functional relationship between a product's sales and independent variables that influence the sales. Several independent variables can be incorporated, including a time index or time-dependent variable. Conditional forecasting is usually performed when the explanatory variables themselves are not known with certainty and therefore must be forecast. Because of the stochastic nature of the predicted values of the independent variables, uncertainty is treated explicitly in such a model.

Econometric models can also consist of more than one equation.[6] In the single-equation regression model, a dependent variable is related to a set of independent variables. However, the model does not explain interdependencies that may exist among the explanatory variables themselves, nor the way in which these explanatory variables are related to other variables. In addition, the single-equation model explains causality in one direction only, i.e., the explanatory variables determine a dependent variable, but there is no "feedback" relationship between the dependent variable and the explanatory variables. Simultaneous and recursive equation systems are the two major types of multiequation models. In a simultaneous equation system, all the endogenous (unknown) variables must be determined simultaneously by solving the equations, whereas in a recursive system the endogenous variables are determined sequentially, one at a time.

One of the major limitations of econometric models is their reliance on historical data, as in the case of time series models. Historical data are not available for very new products. Also, the simpler versions of econometric models may lack explanatory value. In general, however, econometric models can be quite sophisticated for forecasting purposes.

If a single-equation econometric model were developed for the heat pump, the dependent variable could be annual shipments of heat pumps, as in the case of the time series model. In contrast to the time series model, several factors affecting the demand for heat pumps could be included as independent variables. These variables could be technology-related, such as (1) the costs associated with the heat pump, e.g., installation, fuel, or maintenance costs, (2) the energy efficiency ratios for heat pumps, which are available from ARI for the 1976-1983 period, and (3) the capital costs of competing systems, such as a heating system -- gas furnace, electric resistance heating, or oil furnace -- plus air conditioning. Another category of variables could possibly be related to general economic conditions, such as (1) consumers' disposable income, (2) the number of housing starts for both single-family and multifamily homes, and (3) relative fuel costs, such as the price ratios of electricity to natural gas and of electricity to fuel oil. Historical data for such variables could be obtained from various sources.

2.6 DIFFUSION MODELS

The diffusion process is concerned with the spread of a new product in the market. The behavioral theory underlying diffusion models is that new-product acceptance involves both adoption and imitation: the new product is first adopted by a few people -- the innovators -- who in turn influence other people to adopt it. The basic diffusion model can be represented as:[4]

$$\frac{dN(t)}{dt} = a[\overline{N}(t) - N(t)] + b\,N(t)\,[\overline{N}(t) - N(t)]$$

where:

$\dfrac{dN(t)}{dt}$ = rate of diffusion at time t,

$N(t)$ = cumulative number of adopters at time t,

$\overline{N}(t)$ = population of potential adopters at time t (ceiling penetration level),

a = coefficient of innovation (constant), and

b = coefficient of imitation (constant).

The differences among various diffusion models stem from the assumptions made concerning the constants a and b, and whether N(t), the market potential, is a constant or varies over time.

The Fourt and Woodlock model, for instance, was developed to predict the penetration levels of a frequently purchased nondurable consumer item.[11] This model assumes that b = 0, i.e., that the diffusion process occurs purely through the innovation effect and that the ceiling penetration level is a constant, $\overline{N}$. Hence, the number of new adopters at time t is proportional only to the potential number of adopters still available, i.e., $\overline{N} - N(t)$. The result is a concave (exponential growth) diffusion curve, with the increments of penetration for equal time periods being proportional to the remaining distance to the ceiling penetration limit.

Another major diffusion model is the Bass model, which was developed for durable consumer items and has also been applied to retail services, industrial technology, and agriculture.[12] This model also assumes that the ceiling penetration is a constant, $\overline{N}$, but covers both the innovative and imitative effects. In terms of the mathematical formulation above, the coefficient of innovation would be a and the coefficient of imitation would be $b\overline{N}$. That is, the innovators are not influenced in their timing of purchase by the number of persons who have already bought the product. As the acceptance process continues, the relative number of innovators will decrease monotonically with time. Imitators, however, will be influenced by the number of previous adopters and will increase relative to the number of innovators over time. The shape of the diffusion curve therefore depends on the relative rates of these two opposite tendencies. When $b\overline{N} > a$, the sales curve starts at 0, grows to a peak, and then declines. When $b\overline{N} < a$, the curve declines continuously.

Another group of diffusion models consists of technological substitution models, which address the process by which a new product substitutes for an established one. These assume that a = 0 in the basic diffusion equation, making them pure imitation-effect models. The earliest such model was developed by Mansfield[13] and revised by Blackman.[14] The extent of substitution was defined in terms of the market share captured by the new technology. Blackman's basic model is given by:

$$\ln \left[m/(L - m) \right] = -\ln \left[(L/No) - 1 \right] + \gamma \left(t - t_1 \right)$$

where:

m = market share captured at time t by a new innovation,

L = upper limit of the market share that the new innovation can capture in the long run,

No = market share captured when $t = t_1$,

γ = constant that governs substitution rate, and

t_1 = year in which the innovation captures a portion of the market.

The equation can be rewritten as:

$$\ln \left[m/(L - m) \right] = C_1 + C_2 \left(t - t_1 \right)$$

where C_1 and C_2 are constants.

To forecast the rate of substitution of the new innovation, one has to estimate L. Also, C_1 and C_2 must to be estimated from historical data on the market shares of the new technology over time. One advantage of the model is that the historical data need not be extensive. It has been demonstrated that predictions can be made when the new technology has captured as little as 2% of the total market.

Several variations of the Mansfield and Blackman models have been developed.[15] Based on a study of several technological substitutions, Fisher and Pry concluded that once a technology has captured about 5% of the market, the probability is high that it will continue to spread until it completely replaces the former technology.[16] The pattern of substitution is assumed to follow an S-shaped (logistic) curve. The model is mathematically represented as:

$$f/(1 - f) = \exp \left[2\delta \left(t - t_o \right) \right]$$

where:

f = fraction of the market substituted,

δ = half of the annual fractional growth in the early years, and

t_o = time at which $f = 1/2$.

The Fisher and Pry approach has been found to be a useful representation of the substitution process in a number of applications.[17,18]

Many other types of technological substitution models have been developed, but these are beyond the scope of this paper. Hunter and Rubenstein[15] have reviewed many of them. Most suffer from certain restrictions: (1) the consideration of substitutions between only two items, (2) the assumption that once the substitution process begins, it will not reverse itself, and (3) the necessity for a certain amount of historical data to be available from the early stages of the substitution.[19] These restrictions may not apply in all instances.

Technological substitution models provide strong support for the hypothesis that technological substitution follows an S-shaped pattern. In fact, most diffusion models use some type of S-shaped growth curve.[20] The reasons for the slow growth rate in the initial phase include a lack of information, supply bottlenecks, and uncertainties surrounding the product, all causing postponement of purchases. In the intermediate phase, the growth rate accelerates because supply bottlenecks have been overcome and more information is available about the product. Also, interaction between adopters (innovators) and nonadopters (imitators) aids the diffusion process. In the final phase, the growth rate again slows because market penetration approaches the saturation, or ceiling, level. Hence, the number of potential customers who have still not tried the product is relatively small.

Usually a function that limits the range of the dependent variable is chosen to represent product diffusion over time. Common functions include the cumulative normal, logistic, lognormal, and Gompertz functions. The Mansfield-Blackman type of technology substitution models use the logistic growth function. All of these have S-shaped cumulative density functions whose slopes equal the growth rates of demand over time. Lakhani has discussed the mathematical properties of the latter three functions and concludes that the Gompertz is the least restrictive of them.[21] Figure 2.1 illustrates a typical S-curve diffusion model.

Two classes of diffusion models exist with respect to assumptions about the ceiling market potential.[22] Most models (such as those of Bass, Fourt and Woodlock, and Fisher and Pry) assume that the market potential is constant over the time frame of the diffusion process; these are static models. Dynamic models assume that the market potential is time-dependent. Mahajan and Peterson have proposed a dynamic model in which the market potential increases over time but the difference between the market potential and the product growth curve (actual cumulative number of adopters) decreases over time.[23]

Diffusion models also differ in their inclusion of marketing strategy variables such as advertising, promotion, and price. Several extensions of the basic diffusion model have incorporated these variables by considering their effects on various parameters of the diffusion process, namely the imitation coefficient b, the innovation coefficient a, and/or the market potential N(t). Mahajan and Muller[24] compare the basic diffusion models and their extensions in terms of the method of incorporating marketing and other variables in these three parameters.
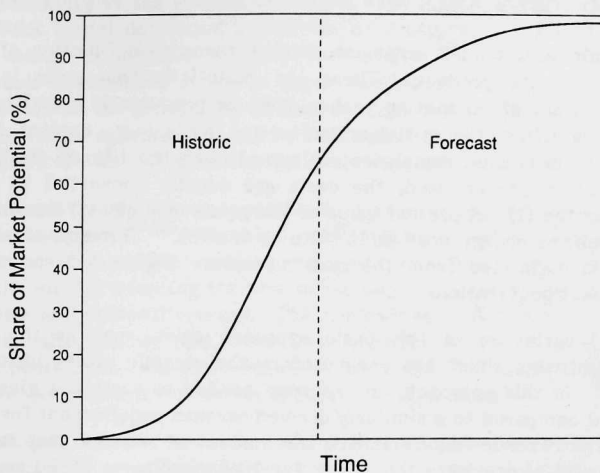
FIGURE 2.1 Typical S-Curve Diffusion Model

Recently, Peterson and Mahajan[25] have extended the static diffusion model to account for certain types of relationships -- independent, complementary, contingent, and substitute -- between the new product and other, existing products in the marketplace. They have also developed methods to incorporate the spatial diffusion pattern of a new product.[26]

To estimate the parameters of the basic diffusion model, it is necessary to have historical data, experimental or market test data, or data on an analogous product. All diffusion models are deterministic in that they provide point estimates of sales. Hence, uncertainty is not treated explicitly. However, one can use different values for the input parameters and a Monte Carlo simulation to develop probabilistic projections. Also, an estimate of the market potential is required. The reliability and validity of diffusion models in relation to other types of models have not been tested adequately. Hence, a global assessment of diffusion models is difficult to make.

If a diffusion model is used to model the market penetration of a new technology, such as the heat pump, the assumption about market potential would be crucial. If a static model is used, the market potential would first have to be estimated. One possible approach would be to estimate the total market for space heating and cooling. Data for this purpose is available from such sources as Data Resources, Inc. (DRI). One must then calculate what proportion of this total market represents the market potential for heat pumps. Actually, the static model would be unrealistic for heat pumps and most durable goods, one reason being that the number of housing starts has tended to change over time. Dynamic diffusion models may therefore be more appropriate (and accurate) for most new technologies.

2.7 ECONOMIC COST MODELS

Economic cost models estimate market share as a function of only the cost-related aspects of the product. There are essentially four steps in economic cost models. First, a set of competing technologies or products is identified. Second, for each product identified, the initial and after-tax costs (e.g., capital, installation, and operation and maintenance costs) are estimated over the life of the product. Third, depending on the approach used, the costs are usually converted to a standard cost measure such as the (1) net present value or life-cycle cost ($), (2) internal rate of return (%), or (3) levelized energy cost ($/10^6 Btu or ¢/kWh).[27] Finally, the market share of each product is estimated from the cost measure. Figure 2.2 shows these steps in estimating market penetration.

Several variations of this basic approach exist, such as the annual revenue requirement approach, which has been used by the electric utility industry to compare alternatives.[28] In this approach, the revenue needed to sustain a given alternative is determined and compared to a similarly derived revenue requirement for each competing alternative. The revenue requirement is the amount of revenue that must be collected from customers to compensate the utility for all expenditures (fixed and variable costs, taxes, and profits) associated with the implementation of an alternative. In the long run, the revenue must pay for all the costs of doing business. Usually, the revenue requirement is calculated for each year over the life of the alternative. The present value is then calculated and levelized over a given time period. Further, for comparison of alternatives, the levelized revenue requirement is often calculated per unit of output. Hence, the measure is essentially the same as the levelized energy cost ($/10^6 Btu or ¢/kWh). The alternative with the least revenue requirement is then selected by the utility.

A modification that is more suited for forecasting market shares of competing products or technologies is to represent the revenue requirement or levelized energy cost of each alternative in terms of a distribution (often normal or triangular) incorporating the uncertainty of the cost estimate. Minimum and maximum costs can be computed using the minimum and maximum ranges expected for the various costs. The ends of each range can be specified as percentages of the expected cost (e.g., 80% to 135%). A probability cost distribution is then constructed using this information. A triangular distribution is simpler than other possible distributions. Such a distribution takes into

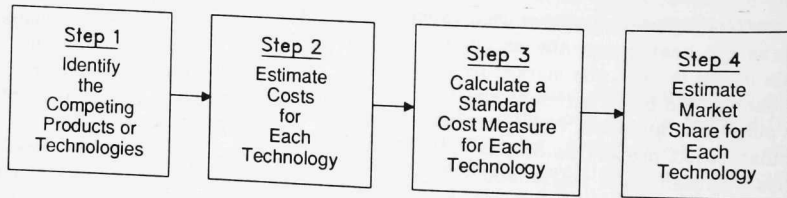| Step 1	Step 2	Step 3	Step 4
Identify the Competing Products or Technologies	Estimate Costs for Each Technology	Calculate a Standard Cost Measure for Each Technology	Estimate Market Share for Each Technology

FIGURE 2.2 Steps in a Typical Economic Cost Model for Estimating the Market Share of Competing Technologies

account the uncertainty in the estimation of the costs. This type of approach has been used in the SPURR model developed by the MITRE Corporation's METREK Division for estimating the market penetration of solar technologies.[29] After a cost distribution is available for each alternative, a Monte Carlo simulation can be used to estimate the market shares based on the probability of each product having the lowest cost or revenue requirement. A similar approach is used by DRI based on the life-cycle cost distribution for each alternative and on simulation to determine the market shares.

Economic cost models have been widely applied to estimate the market penetration of solar technologies.[29,30] In these models, market penetration is based on the S-shaped logistic curve. The vertical axis measures the market share, i.e., the percentage of the market adopting the new technology. The horizontal axis measures the technology's economic competitiveness. The methods and criteria used to compare the economics of the alternatives vary. The least complex method is to compare payback periods. Other possibilities are comparisons of life-cycle cost or of the number of years to reach positive savings, and various combinations of measures.

The S-shaped logistic curve (see Fig. 2.3) represents the theory that, when the solar technology is only marginally better than the conventional technology, only a few innovators will adopt it. As the solar technology becomes economically superior, its market share increases rapidly but the growth rate again slows when most of the market has been captured.[30] A major criticism of the S-shaped curve is that little empirical information exists to support such a relationship between economic attractiveness and market share.[29,30] The S-shaped logistic curve has been borrowed from diffusion models, where it is used to model market share over time. Its validity is in question in the case of economic cost models.

In the case of the heat pump, the major competing technologies would probably be electric resistance heating, natural gas and fuel oil furnaces, and air conditioning systems. Cost information on these technologies could be collected fairly easily. A triangular distribution for levelized energy cost could then be assumed, and simulation techniques could be used to calculate the market share for each technology.

The advantage of such a model is that cost factors can be specified in a relatively straightforward manner. In addition, sensitivity analysis can be performed on the cost factors. The main disadvantage is that noncost aspects are ignored. In addition, the market share calculated from the model may not be attained immediately. It might take some time in spite of the economic attractiveness of the technology. This delay is not incorporated in the economic cost model per se and therefore a combination of the economic cost and diffusion models might be needed. Such a combination is discussed in Sec. 3.

2.8 DISCRETE CHOICE MODELS

Discrete choice models predict the probability of purchasing a new or existing product. The probability of purchase is a function of the utility of the product, and it is assumed that the consumer will choose the product that provides the highest utility. Discrete choice models apply to situations where choices are made among separate,
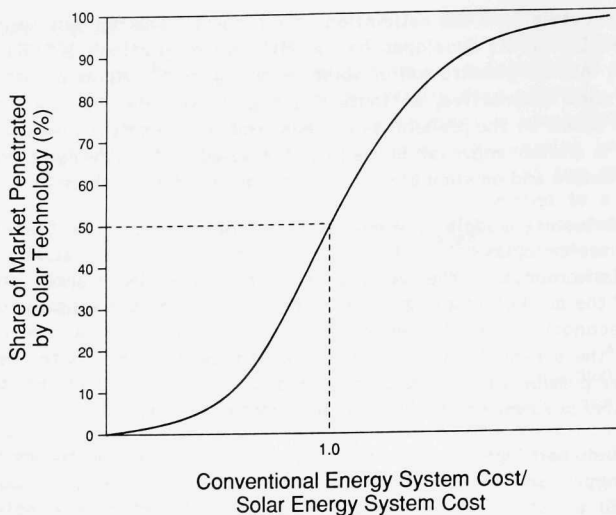
FIGURE 2.3 Use of an S-Shaped Logistic Curve to Estimate
Market Penetration

distinct alternatives.   In other words, one cannot choose a partial amount of each
alternative. Such models have been used in a number of applications such as choices of
energy alternatives (e.g., space heating systems) transportation mode, or durable
consumer items.

The utility of the product is usually specified as a function of product
characteristics (e.g., cost factors, efficiency) and consumer characteristics (e.g., age,
income, education, or race).   The probability of adopting a new technology is thus
estimated from several independent variables characterizing the technology and the
individual. Assuming that information on these attributes is available for a sample of
individuals, an equation can be estimated that will predict the choices of individuals not
in the original sample.[31]   Discrete choice models thus enable such statements as the
following to be made: "The probability is 0.4 that, in the South, a new-home buyer with
an income of $40,000 will opt for a heat pump."

Discrete choice models can generally be classified into two categories: binary-
choice and multiple-choice models.[6] Binary-choice models are used for a choice between
two alternatives, e.g., adoption or nonadoption of the home heat pump. The dependent
variable in this case would be a dichotomous random variable that assumes a value of 1 if
the heat pump is adopted and a value of 0 otherwise. The probability model can be
specified such that the probability of adoption of the heat pump is a function of
consumer and technology characteristics. The most commonly used binary-choice models
are the linear probability, probit, and logit models. Some problems exist with the linear

probability models, most notably that the probability predictions may lie outside the (0,1) interval. Hence, the probit and logit models are preferred. The probit model is associated with the cumulative normal probability function, and the logit model with the cumulative logistic probability function.

When the choice involves three or more alternatives, multiple-choice models are used. The choices are assumed to be mutually exclusive. The multinomial logit and multinomial probit models are commonly used. The Gas Research Institute has used a multinomial logit model to study household demand for home space heating systems and energy inputs, household water-heating fuel choices, and household adoption of washers and dryers.[32] Factors such as the perceived benefits of the option relative to others, energy price differences, climatic conditions, housing unit conditions, and household-specific characteristics are used to predict relative adoption probabilities.

Burton, Jackson, and Kaserman describe a discrete choice model for space heating and cooling system combinations.[33] They suggest that the market penetration of new technologies is determined primarily by (1) growth of the capital stock, (2) replacement of the existing stock, and (3) purchase choice of the new technology relative to the existing technology on the basis of the attributes of alternative technologies. This discrete choice model provides the latter item, and the first two items are incorporated into the model through a simulation framework. Hence, the model is truly a combination of the discrete choice model and an equipment stock model. The authors used the model to estimate the market penetration of gas heat pumps in Kansas over a 30-yr period. The sequential logit model (a variation of the multinomial logit model, which structures the decision process as a sequence of decisions) was used to predict purchase probabilities, while replacement decisions were simulated by a relationship between the variable cost of existing equipment and the total cost of new equipment. Stock growth was assumed to be represented by historical growth. The authors conclude that the forecast matches other historically observed penetration functions. They state that their "simulation exercise indicates that it is possible empirically to model the complex interaction between equipment stock growth and purchase choice in a manner consistent with economic theory and engineering characteristics of competing systems."

Discrete choice models can be used to make market share predictions. The probability of adopting a technology is modeled as a function of individual or household characteristics. Hence, if the number of households with a particular combination of attributes is known, the expected probability of adoption of the technology for the entire population, i.e., the technology's market share, can be calculated. One difficulty, however, is the "aggregation problem." The model uses individual household data to estimate the purchase probability for each household. Aggregating the purchase probabilities of individual households to predict market share for the whole population is not statistically correct. One approach to solving this problem is to use average population values for the independent variables so that the purchase probability for an "average" household can be determined. The preferred approach, however, is to first divide the households into homogeneous clusters. Then, the average values of the independent variables for each cluster are used to determine the average purchase probability for each cluster. The weighted average of these cluster probabilities, with the weight being the cluster size, gives the aggregate population purchase probability.

3 COMBINED METHODOLOGIES

Section 2 discussed several approaches to estimating the market penetration for new technologies. Often, however, a single approach may be insufficient. A combination of different approaches may provide better explanatory or predictive power by incorporating the advantages of each approach. In this section, two such combination approaches to estimating market penetration are considered. This discussion by no means exhausts all possible combinations; it merely illustrates that different methodologies can complement one another.

3.1 ECONOMIC COST MODEL AND DIFFUSION MODEL

Economic cost and diffusion models have been combined for solar technology applications. The solar penetration model developed by SRI[29] is used to estimate market penetration for every 5-yr period from 1975 to 2020. In this model, solar technologies that can compete in various end-use markets are compared, on a cost-per-million-Btu basis, with the conventional energy sources for each sector. The pertinent end-use markets are defined by the type of demand (e.g., water heating, space heating, space cooling, process heat, and on-site, base, intermediate, and peak electricity) and the end-use sector (i.e., residential/commercial, industrial, and electric utility). These cost comparisons are the basis for the market penetration estimates. For each end-use market in which a solar technology can compete, a ratio of the price of energy delivered from the solar system to the average price of conventional energy is calculated. A gamma function is used to relate this price ratio to the "idealized" market share of the solar technology. For example, when two energy sources produce energy that is equal in price, each receives 50% of the market under equilibrium conditions. A "gamma parameter" is subjectively specified for each technology to account for market imperfections such as price variations, consumer preferences, and consumer resistance to change. This gamma function represents the cost-model portion of the SRI model.

The diffusion model is incorporated by defining a behavioral lag function to account for the time required for a new technology to achieve its full market potential. The behavioral lag function is defined as the time required for half of the market to respond to the new technology; this is subjectively estimated. The share of a particular market captured by a new technology is determined as follows:

$$\text{Market Share} = \begin{bmatrix} \text{Idealized Market} \\ \text{Share (from} \\ \text{cost model)} \end{bmatrix} \times \begin{bmatrix} \text{\% of Market Responding} \\ \text{within a Certain Time} \\ \text{(from diffusion model)} \end{bmatrix}$$

A method similar to the SRI approach has been applied to new energy systems by Teotia et al.[34] The model has four phases:

1. Segmenting the new technology and/or end-use market,

2. Estimating the technical market,

3. Estimating the economic market, and

4. Estimating the degree of market penetration.

In Phase 1, segmentation is based on the decision makers' objectives and resource constraints. Segmentation is possible by system size, end-use sector, or region. In Phase 2, the technical market includes all feasible commercial applications of a new technology or system. To estimate the current technical market, data are collected for all applications in which conventional systems can be replaced by new systems. Econometric models based on trends in the end-use markets are used to forecast the markets for conventional systems.

In Phase 3, the economic market is the portion of the technical market in which a new energy system is economically more attractive than its competing systems. This is the phase in which the economic cost model is applied. As described in Sec. 2.7, this model estimates the probability of a particular new energy system being the most economical. The economic market for the new technology is then determined as the technical market multiplied by the probability of the system being most economical.

Finally, in Phase 4, the diffusion model is applied. New energy systems cannot be expected to capture the economic market immediately. Penetration of the economic market over time can be estimated with an S-shaped curve such as the Gompertz curve. Teotia et al. use the behavioral lag function designed by SRI.

The methodology described above has been applied successfully by Teotia et al. to a number of cases, including cogeneration energy systems[34-36] and new energy-efficient electric motors.[37]

3.2 DISCRETE CHOICE MODEL AND DIFFUSION MODEL

When the discrete choice and diffusion models are combined, the economic market for the technology is defined as a function of two factors: the total potential market for the technology and the probability of choosing the new technology. The total potential market for a new technology can be estimated by an equipment stock model or by trends within the conventional technology market. The probability of choosing the new technology over the conventional technology can be estimated with a discrete choice model. The economic market for the new technology can then be estimated by multiplying the product of the total potential market by the probability of choosing the new technology.

Again, the economic market may not be attained immediately, but only gradually over time. This process can be modeled with a diffusion curve. The combination of the discrete choice and diffusion models is thus very similar to that of the economic cost and diffusion models. However, no empirical application of this methodology is available in the literature.

4 CONSIDERATIONS IN METHODOLOGY SELECTION

This section describes certain general considerations in selecting a market penetration forecasting methodology for a particular situation.

In business and industry, the simple methods of forecasting are used most often. In one study, where 20 interviews were conducted in six industries, nearly all respondents used expert judgment or simple time series models to forecast product sales.[27] Reliance on these simple methods may be attributed to several factors, including a lack of awareness of many sophisticated methods, restrictions of the available data, and skepticism about the assumptions underlying various other methods. However, less sophisticated methods are not necessarily inferior. In fact, less expensive studies conducted recently may be more accurate than older, more sophisticated studies because the core assumptions may be outdated in the latter.[28] Hence, one must consider the entire range of methodologies, from the simplest to the most sophisticated, when selecting a methodology for a particular situation. Also, combinations of methodologies may be superior to any one methodology, which further increases the options available.

The process of selecting a methodology for a forecasting situation is an iterative one. The choice must first be narrowed to a few acceptable methodologies; then, one or a combination of methodologies can be selected as the most appropriate. In general, the process consists of three steps. The first step is to clearly identify the objectives of the forecasting task, the expected results, and the way in which the results will be used. The second step is to characterize the technology and its market. Generally, such characteristics as the size of the current market, historical and anticipated growth in the market, and interproduct competition should be considered in this step. The third step is to select the methodology. The criteria include the risk of the technology, the accuracy desired in the forecasting, the stage in the life cycle of the technology, the extent and nature of available data, and the time horizon of the forecast. Each of these factors is briefly discussed below.

The first factor is the risk of the technology, i.e., the potential impact of the technology in the industry and the uncertainty surrounding this impact. Although generalization is difficult, simpler methods such as historical analogy and time series may be more appropriate in low-risk situations. Econometric, diffusion, cost, and discrete choice models, as well as consumer surveys, are desirable for high-risk technologies.

The forecasting accuracy of the various methodologies is difficult to assess. Forecasters often use a combination of methods, even though one is designated as the primary methodology. Thus, the accuracy of a methodology in isolation may be in question. Moreover, the accuracy of forecasts is only loosely tied to methodological sophistication. In many cases, simple methods do as well as, if not better than, sophisticated methods. One factor, however, that is definitely linked to accuracy is the

validity of the core assumptions underlying the methodology, as the following excerpt from an article by W. Ascher describes:

> The core assumptions are the major determinants of forecast accuracy. When the core assumptions are valid, the choice of methodology is either secondary or obvious. A sophisticated methodology cannot redeem a forecast based on faulty core assumptions.[38]

For example, diffusion models assume an underlying behavioral process in the market growth of a product. If this process is not valid for a product, use of such a model is bound to result in an inaccurate forecast. Another aspect related to the accuracy of a forecast is the consideration of uncertainty. In many instances, it is desirable to use a methodology that explicitly incorporates uncertainty. If not, sensitivity analyses on the explanatory or causal variables may be needed. Confidence intervals are generally more informative than point estimates. Econometric, economic cost, and discrete choice models can incorporate uncertainty, whereas diffusion models cannot.

The stage in the life cycle of a product or technology is a third factor in the choice of a methodology. In the preintroductory stage, little information is available, and historical analogy or consumer surveys (primary data) may be the only choices available. In the earlier stages of the life cycle, cost models and diffusion models become viable alternatives. Time series, econometric models, and discrete choice models become relevant later in the life cycle as these methodologies require more data.

The extent and nature of available data may also play a role in methodology selection. The amount of data required is largely a function of the objectives and needs of the particular application. However, certain methodologies may be better suited for handling certain types of data; for example, time series models are excellent at handling cyclical patterns.

The time horizon of the forecast must also be considered. Short-range forecasts may have different requirements than do long-range forecasts. Short-range forecasts usually must be more accurate and must use methodologies that yield quick results. Use of time series models with smoothing methods is popular for short-term forecasts. Time series with ARMA methods are more suited for the medium term, whereas econometric models and diffusion models are more suited to the medium to long term.

Other factors that should be considered in the choice of a forecasting methodology include the resources available for the task, ease of application by the user, and expertise of the model builder. The implications of these factors are fairly clear and will not be discussed here. Ultimately, however, methodology selection reduces to a comparison of cost and value. The cost may include the dollar cost as well as the time and effort involved in the forecast. The value of a methodology chiefly depends on the accuracy of the forecast yielded and the importance of having an accurate forecast. While any forecast will involve some error, the accuracy of the forecast typically increases with the sophistication of the methodology (however, the cost of the forecasting also increases). Normally, users attempt to choose the methodology that gives the best value for the least cost.

5 CONCLUSION AND FUTURE WORK

Several methodologies are presently available for forecasting the market penetration of new technologies. This report has reviewed the major methodologies, which range from simple methods such as subjective estimation to more-complex ones such as diffusion models and discrete choice models. To choose a market penetration methodology for a particular situation, one must know the range of methodologies available and the essential characteristics, advantages, and limitations of each. This report has attempted to provide this information along with examples to show how a particular methodology has been or can be used in various applications.

A case example of selecting a methodology for forecasting demand of advanced electric heat pumps in residential sectors will be provided in a forthcoming report. Each of the methodologies surveyed in this report will also be considered in the forthcoming report, specifically in terms of their suitability for projecting the heat pump market by selected market segments, including minority households. Based on the methodology selected, the report will provide a heat pump market penetration model, and describe its application for projecting the demand for heat pumps in such key market segments as type of home (single-family or multifamily), type of household (white, black, or Hispanic), and type of construction (new or old).

REFERENCES

1. Claycamp, H.J., and L.E. Liddy, *Prediction of New Product Performance: An Analytical Approach*, J. of Marketing Research, 6:414-420 (Nov. 1969).

2. Urban, G.L., *SPRINTER MOD III: A Model for the Analysis of New Frequently Purchased Consumer Products*, Operations Research, 18:805-855 (Sept. 1970).

3. Wheelwright, S., and S. Makridakis, *Forecasting Methods for Management*, John Wiley, New York (1980).

4. Wind, Y.J., *Product Policy: Concepts, Methods, and Strategy*, Addison-Wesley (1982).

5. *Heat Pump Technology*, Gordian Associates Inc., prepared for the U.S. Department of Energy, Contract EX-76-C-01-2121, p. 419 (1978).

6. Pindyck, R.S., and D.L. Rubinfeld, *Econometric Models and Economic Forecasts*, McGraw-Hill (1976).

7. Fitzroy, P.T., *Analytical Methods for Marketing Management*, McGraw-Hill (1976).

8. Croxton, F.E., and D.J. Cowden, *Applied General Statistics*, Prentice-Hall, Englewood Cliffs, N.J. (1955).

9. Box, G.E.P., and G.M. Jenkins, *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco (1970).

10. *Statistics*, Air Conditioning and Refrigeration Institute, Arlington, Va. (1984).

11. Fourt, L.A., and J.W. Woodlock, *Early Prediction of Market Success for New Grocery Products*, J. of Marketing, 25:31-38 (Oct. 1960).

12. Bass, F.M., *A New Product Growth Model for Consumer Products*, Management Science, 15(5):215-227 (Jan. 1969).

13. Mansfield, E., *Technical Change and the Rate of Imitation*, Econometrica, Vol. 29 (Oct. 1961).

14. Blackman, A.W., *The Market Dynamics of Technological Substitutions*, Technological Forecasting and Social Change, 6:41-63 (1974).

15. Hunter, A.P., and A.H. Rubenstein, *Market Penetration by New Innovations: The Technological Literature*, Technological Forecasting and Social Change, 11:197-221 (1978).

16. Fisher, J.C., and R.H. Pry, *A Simple Substitution Model of Technological Change*, Technological Forecasting and Social Change, 3:75-88 (1971).

17. Lenz, R.C., and H.W. Lanford, *The Substitution Phenomenon*, Business Horizons, 15(1):63-68 (1972).

18. Twiss, B.C., *Forecasting Market Size and Market Growth Rates for New Products*, J. of Product Innovation Management, 1(1):19-29 (Jan. 1984).

19. Stover, J.G., *Use of Decision Modeling for Substitution Analysis: Application to Acceptance of New Electricity Generating Technologies*, Technological Forecasting and Social Change, 12:337-351 (1978).

20. Rogers, E.M., and F.F. Shoemaker, *Communication of Innovations: A Cross-Cultural Approach*, 2nd Ed., Free Press, New York (1971).

21. Lakhani, H., *Empirical Implications of Mathematical Functions Used to Analyze Market Penetration of New Products: Cigarettes Case Study*, Technological Forecasting and Social Change, 15:147-156 (1979).

22. Mahajan, V., and R.A. Peterson, *First-Purchase Diffusion Models of New-Product Acceptance*, Technological Forecasting and Social Change, 15:127-146 (1979).

23. Mahajan, V., and R.A. Peterson, *Innovation Diffusion in a Dynamic Potential Adopter Population*, Management Science, 15:1589-1597 (Nov. 1978).

24. Mahajan, V., and E. Muller, *Innovation, Diffusion and New Product Growth Models*, J. of Marketing, 43:55-58 (Fall 1979).

25. Peterson, R.A., and V. Mahajan, *Multi-Product Growth Models*, in Research in Marketing, J.N. Sheth, ed., JAI Press, Greenwich, Conn. (1978).

26. Mahajan, V., and R.A. Peterson, *Integrating Time and Space in Technological Substitution Models*, Technological Forecasting and Social Change, 14 (Aug. 1980).

27. *Methods for Analyzing the Market Penetration of End-Use Technologies: A Guide for Utility Planners*, Electric Power Research Institute Report, Palo Alto, Calif., EA-2702 (1982).

28. *Technical Assessment Guide*, Electric Power Research Institute, Palo Alto, Calif., Report PS-1201-SR (July 1979).

29. Schiffel, D., et al., *The Market Penetration of Solar Energy: A Model Review Workshop Summary*, Solar Energy Research Institute, Golden, Colo. (Jan. 1978).

30. Warren, E.H., Jr., *Solar Energy Market Penetration Models: Science or Number Mysticism?*, Technological Forecasting and Social Change, 16:105-118 (1980).

31. McFadden, D., *Conditional Logit Analysis of Qualitative Choice Behavior*, in Frontiers in Econometrics, P. Zerembka, ed., Academic Press, New York, pp. 105-142 (1974).

32. Tolley, G.S., R.J. Krumm, and E.K. Mensah, *Market Penetration Models and the Evaluation of Gas Technology: Final Report*, Gas Research Institute, Chicago (July 1982).

33. Burton, E., J.R. Jackson, and D.L. Kaserman, *The Discrete Choice Model for New Technologies and Its Application in Modeling Market Penetration*, Electric Power Research Institute, Palo Alto, Calif., Report EA-1961, pp. 431-436 (Aug. 1981).

34. Teotia, A.P.S., C. Lee, and A.S. Kennedy, *Market Penetration of New Energy Systems Estimated by Econometric and Stochastic Methodology*, Proc. Institute of Gas Technology Energy Modeling II Symp., Colorado Springs, Colo. (Aug. 1979).

35. Teotia, A.P.S., A.A. Davis, and D. Poyer, *Cogeneration Energy Systems Show Promising Market Potential in Residential and Commercial Sector over Next Twenty-five Years*, Proc. 16th Intersociety Energy Conversion Engineering Conf., American Society of Mechanical Engineers, Atlanta, Ga. (Aug. 1981).

36. Teotia, A.P.S., and D.J. Bingaman, *Estimating Impact of Changing Energy Scenarios on New Energy Systems: A Case Example of District Heating Cogeneration Systems*, Proc. Energy, Power, and Environmental Systems Conf., sponsored by the International Assn. of Science and Technology for Development, San Francisco (June 1984).

37. Teotia, A.P.S., Y.L. Klein, and F.R. Wyant, *Estimating Market Penetration of New Energy Products: A Case Example of New Energy-Efficient Electric Motors, Modeling and Simulation*, Vol. 12, Proc. 12th Annual University of Pittsburgh Modeling and Simulation Conf., Pittsburgh (May 1981).

38. Ascher, W., *Problems of Forecasting and Technology Assessment*, Technological Forecasting and Social Change, 13:149-156 (1979).